Docket No. NEB-232

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
APPLICATION FOR UNITED STATES LETTERS PATENT

INVENTOR(S): Keith D. Lunnen
Theodore Davis
Geoffrey G. Wilson

TITLE: Method For Cloning And Expression of *Sbf*I Restriction Endonuclease And *Sbf*I Methylase In *E. coli*

ATTORNEY: Harriet M. Strimpel
NEW ENGLAND BIOLABS INC.,
32 Tozer Road
Beverly, Massachusetts 01915
(978) 927-5054; Ext. 373

Customer No.: 28986

Express Mailing Label No.: ER279330987US

METHOD FOR CLONING AND EXPRESSION OF *SBFI* RESTRICTION ENDONUCLEASE AND *SBFI* METHYLASE IN *E.COLI*

BACKGROUND OF THE INVENTION

Present embodiments of the invention relate to recombinant DNA that encodes the SbfI restriction endonuclease (SbfI endonuclease; R.SbfI) as well as the SbfI modification methyltransferase (SbfI methyltransferase; M.SbfI), and to the expression of the SbfI endonuclease and methyltransferase in *E. coli* cells that contain the recombinant DNA.

Restriction endonucleases are enzymes that occur naturally in certain unicellular microbes—mainly bacteria and archaea—and that function to protect these organisms from infections by viruses and other parasitic DNA elements. Restriction endonucleases bind to specific sequences of nucleotides ('recognition sequence') in double-stranded DNA molecules (dsDNA) and cleave the DNA, usually within or close to the sequence, disrupting the DNA and triggering its destruction. Restriction endonucleases commonly occur with one or more companion enzymes termed modification methyltransferases. Methyltransferases bind to the same sequences in dsDNA as the restriction endonucleases they accompany, but instead of cleaving the DNA, they alter it by the addition of a methyl group to one of the bases within the sequence. This methylation ('modification') prevents the restriction endonuclease from binding to that site thereafter, rendering the site resistant to

cleavage. Methyltransferases function as cellular antidotes to the restriction endonucleases they accompany, protecting the cell's own DNA from destruction by its restriction endonucleases. Together, a restriction endonuclease and its companion modification methyltransferase(s) form a restriction-modification (R-M) system, an enzymatic partnership that accomplishes for microbes what the immune system accomplishes, in some respects, for multicellular organisms.

A large and varied class of restriction endonucleases have been classified as 'Type II' restriction endonucleases. These enzymes cleave DNA at defined positions, and in purified form can be used to cut DNA molecules into precise fragments for gene cloning and analysis. The biochemical precision of Type II restriction endonucleases far exceeds anything achievable by chemical methods, making these enzymes the reagents sine qua non of molecular biology laboratories. In this capacity, as molecular tools for gene dissection, Type II restriction endonucleases have had a profound impact on the life sciences in the past 25 years, transforming the academic and commercial arenas, alike. Their utility has spurred a continuous search for new restriction endonucleases, and a large number have been found. Today more than 200 Type II endonucleases are known, each possessing different DNA cleavage characteristics (Roberts and Macelis, Nucl. Acids Res.29:268-269 (2001)). (REBASE®, http://rebase.neb.com/rebase). Concomittantly, the production and purification of these enzymes has been improved by the

cloning and over-expression of the genes that encode them in non-natural production strain host cells such as *E. coli*.

Since the various restriction enzymes appear to perform similar biological roles, in much the same ways, it might be thought that they would resemble one another closely in amino acid sequence and behavior. Experience shows this not to be true, however. Surprisingly, far from resembling one another, most Type II restriction enzymes appear unique, resembling neither other restriction enzymes nor any other known kind of protein. Type II restriction endonucleases seem to have arisen independently of one another for the most part during evolution, and to have done so hundreds of times, so that today's enzymes represent a heterogeneous collection rather than a discrete family. Some restriction endonucleases act as homodimers, some as monomers, others as heterodimers. Some bind symmetric sequences, others asymmetric sequences; some bind continuous sequences, others discontinuous sequences; some bind unique sequences, others multiple sequences. Some are accompanied by a single methyltransferase, others by two, and yet others by none at all. When two methyltransferases are present, sometimes they are separate proteins, at other times they are fused. The orders and orientations of restriction and modification genes vary, with all possible organizations occurring. Several kinds of methyltransferases exist, some methylating adenines (m6A-MMases), others methylating cytosines at the N-4 position (m4C-MMases), or at the 5 position (m5C-MMases). Usually there is no way of predicting, a *priori*,

which modifications will block a particular restriction endonuclease, which kind(s) of methyltransferases(s) will accompany that restriction endonuclease in any specific instance, nor what their gene orders or orientations will be.

From the point of view of cloning a Type II restriction endonuclease, the great variability that exists among restriction-modification systems means that, for experimental purposes, each is unique. Each enzyme is unique in amino acid sequence and catalytic behavior; each occurs in unique enzymatic association, adapted to unique microbial circumstances; and each presents the experimenter with a unique challenge. Sometimes a restriction endonuclease can be cloned and over-expressed in a straightforward manner but more often than not it cannot, and what works well for one enzyme can work not at all for the next. Success with one is no guarantee of success with another.

SUMMARY OF THE INVENTION

In an embodiment of the invention, an isolated DNA encoding the *Sbf*I restriction endonuclease is provided where the isolated DNA is obtainable from *Streptomyces species Bf-61*.

In an additional embodiment of the invention, a recombinant DNA vector is provided that includes a vector into which a DNA segment encoding the *Sbf*I restriction endonuclease

has been inserted. A host cell transformed by the recombinant vector is further provided.

In an additional embodiment of the invention, an isolated DNA encoding the *Sbf*I restriction endonuclease and *Sbf*I methylase is provided where the isolated DNA is obtainable from ATCC No. PTA-5371. A vector that includes this isolated DNA and a host cell transformed by the vector is further provided.

In an additional embodiment of the invention, a method is provided for producing recombinant *Sbf*I restriction endonuclease that includes culturing a host cell transformed with any of the vectors described above under conditions suitable for expression of the endonuclease.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1. Gene organization of the *Sbf*I R-M system. *Sbf*IR, *Sbf*I restriction endonuclease gene; *Sbf*IM, *Sbf*I methylase gene.

Figure 2. The *Sbf*I methylase gene sequence (*SbfIM*, 1460 bp) (SEQ ID NO:1) and the encoded amino acid sequence (SEQ ID NO:2).

Figure 3. The *Sbf*I endonuclease gene sequence (*SbfIR*, 971 bp) (SEQ ID NO:3) and the encoded amino acid sequence (SEQ ID NO:4).

Figure 4. A plasmid map of pACYC184-PstIM clone.

Figure 5. A plasmid map of pACYC184-SbfIM clone.

Figure 6. A plasmid map of pCAB16.

Figure 7. A plasmid map of pLT7K-SbfIR endonuclease clone.

Figure 8. Recombinant *Sbf*I endonuclease activity in cell extract. l DNA was used as the substrate. Lanes 2-8, 1x, 1/2, 1/4, 1/8,1/16,1/32,1/64 diluted cell extract added in the restriction digestions. Lane 1, l DNA digested with native *Sbf*I.

DETAILED DESCRIPTION OF THE INVENTION

The SbfI endonuclease and methyltransferase are enzymes that occur in the bacterium *Streptomyces* species Bf-61 (New England Biolabs' strain collection). The SbfI endonuclease binds to the symmetric nucleotide (nt) sequence 5'-CCTGCAGG-3' in double-stranded DNA molecules (dsDNA) and cleaves the DNA between the A and G in each strand thus: 5'-CCTGCA/GG-3', producing DNA fragments with 4-nt cohesive ends (/ indicates the position of strand-cleavage). Many restriction endonucleases that occur in nature are accompanied by protective modification methyltransferases. However, restriction endonucleases that recognize and cleave long, infrequently occurring sequences

such as R.SbfI, are not always accompanied by a protective methyltransferase. At the outset of these experiments it was not known whether a modification methyltransferase in fact accompanied R.SbfI.

Hurdles in Cloning SbfI restriction endonuclease

(a) Unsuccessful methylase selection

The methylase selection method described in U.S. Patent No. 5,200,333 is the preferred method for cloning restriction-modification systems. Methylase selection did not however yield the *SbfI* methylase gene (*M.SbfI*). The reasons for failure could include any or all of the following technical difficulties: construction of initial libraries where potential cloning sites may cut within the methylase gene; failure to clone the proper fragment from the libraries due to size of the DNA fragment; gene toxicity; low expression of the methylase gene, or the complete absence of the accompanying methylase gene, as with certain other 8-nucleotide specific restriction endonucleases such as *Pac*I 5-TTAAT/TAA-3' (U.S. Patent No. 5,098,839).

(b) Absence of SbfIM downstream of SbfIR

Although the possibility existed that there was no M.SbfI, alternative approaches to cloning the gene were attempted. Subsequently an SbfI methyltransferase was indeed identified by a PCR reaction on the original pUC19-*Sau*3AI partial *Sbf*I DNA library that failed to yield a selectable *sbfIM* gene. The PCR reaction showed the presence of a *Sau*3AI partial fragment that

contained the *sbfIM* region. The location of the *sbfIM* region on chromosomal DNA was eventually determined by inverse PCR to be not downstream of the *sbfIR* gene as first thought but upstream of the *sbfIR* gene. The sequence of *sbfIM* was subsequently obtained.

The *Sbf*I methylase gene was cloned by PCR from *S. Bf-61* chromosomal DNA into pACYC184 and transformed into *E. coli.* The cloned SbfI methyltransferase bound to the same nucleotide (nt) sequence in dsDNA as R.SbfI, and catalyzed the addition of a methyl group to the adenine residue in each strand (5'-CCTGCmAGG-3'), producing modified DNA molecules that was resistant to cleavage by R.SbfI (mA indicates the modified base).

The *Sbf*I restriction endonuclease gene (*sbfIR*) was identified using PCR primers based on the *Sbf*I endonuclease (R.*Sbf*I) amino acid sequence from the N-terminal end and the cyanogen bromide-digested ends of the protein. The gene was subsequently cloned by inverse PCR from *Sbf*I chromosomal DNA.

(c) Inconclusive Expression Levels for Endonuclease

The PCR fragment obtained by inverse PCR containing the *sbfIR* gene was inserted into the plasmid vectors pRRS and pLT7K and used to transform M.*Pst*I pre-modified *E. coli.* M.PstI premodification was used because M.SbfI had not yet been identified and *Pst*I methylase (M.*Pst*I) was shown to protect *Sbf*I sites in addition to PstI sites against both *Pst*I and *Sbf*I endonuclease digestion of purified PstI methylated DNA.

No clones with inserts were found with pRRS; some clones were found with pLT7K where the regulated T7 expression vector pLT7K contained a constitutive anti-sense promoter downstream of *sbfIR* to reduce basal expression. Some pLT7K clones contained correct DNA sequence, however these produced very low *Sbf*I endonuclease activity upon induction. This negative result suggested there was selection pressure to isolate endonuclease mutants with reduced activity. This disappointing result suggested that highly expressing clones were being selected against perhaps because of under-methylation of the host chromosome DNA by M.*Pst*I.

Since expression from a medium-copy-number T7 vector in *E. coli* pre-modified with M.*Pst*I did not generate a stable high expression clone, efforts were made to express the *sbfIR* gene in M.*Sbf*I methylated *E.coli* using pLT7K. When the *Sbf*I endonuclease gene was cloned in M.*Sbf*I pre-modified *E. coli*, a stable and over-expressing clone was established. Over-expression of an enzyme is generally intended to mean at least 10^5ug including 10^6 or 10^7ug. Low expression is less than 10^3ug. Low expression levels of a putative cloned restriction endonuclease may result in cleavage profiles of DNA. However, this does not conclusively prove that the desired enzyme has been obtained. For example, the enzyme digest may be partial or incomplete making it unclear whether the products are merely the result of random cleavage.

(d) Obtaining an over-expressing clone

A stable over-expressing clone of R.SbfI was obtained as follows: The *sbfIR* gene was amplified by PCR from genomic DNA. Following purification, the resulting PCR fragment was blunt-end ligated into pCAB16 at a *Bsa*AI site. pCAB16 is a pUC18 derivative containing the *mspIR* gene in the polylinker of pUC18 in line with the Plac promoter. pCAB16 contains a single *Bsa*AI site within the *mspIR* gene. Insertions at this site interrupt *mspIR* expression (which would otherwise be lethal) enabling plasmids containing inserts to be selectively recovered with high efficiency (Figure 6). The *sbfIR* PCR-fragment was ligated into the *Bsa*AI site of pCAB16, and transformed into M.*Pst*I pre-modified *E. coli*. Clones that were found to carry the *sbfIR* PCR insert were cultured. However, assays showed these clones had no detectable *Sbf*I endonuclease activity. DNA sequencing of these clones showed that an intact *sbfIR* gene was present in the opposite orientation to Plac and *mspIR*, which could explain the lack of R.*Sbf*I activity.

Taking an alternative approach, the *sbfIR* gene was purified from the pCAB16-SbfIR plasmid by gel purification. The resulting DNA fragment was ligated into pLT7K and transformed into M.*Pst*I pre-modified *E. coli*. Clones found to carry the PCR insert were induced with IPTG and assayed for *Sbf*I activity on λ DNA. The extracts generated partial *Sbf*I digestion pattern. DNA sequence of one of these clones showed that it carried an intact *sbfIR* gene. Since the DNA sequence was correct for this pLT7K-*sbfIR* plasmid, most likely highly expressing clones were being

selected against during induction. When the *sbfIM* gene was subsequently identified and cloned into pACYC184, this same pLT7K-SbfIR clone was then transformed into M.*Sbf*I pre-modified *E. coli*. Transformants in which the *sbfIM* gene was expressed from a low-copy-number plasmid pACYC184 and the *sbfIR* gene was expressed using the medium-copy-number pLT7K within the same *E. coli* host were cultured and their cells extracts were assayed for *Sbf*I activity on λ DNA. The recombinant *Sbf*I endonuclease yield was $\sim 10^5$ units/g of wet cells from the over-producing strain.

In summary, an expression strategy was ultimately developed which overcame a number of hurdles and ultimately proved successful in yielding over-expressed R.SbfI. This strategy relied in one embodiment on expressing the R.SbfI and M.SbfI under different strength promoters namely a medium copy number promoter and a low copy number promoter respectively. However, in alternative embodiments, both sbfIM and sbfIR genes may be expressed under the same promoter.

Moreover, the *sbfIM* gene may be cloned in a single plasmid together with the *sbfIR* gene under the same or different promoters or in seperate plasmids under the same or different promoters.

The method described herein by which the *sbfIM* and *sbfIR* genes are preferably cloned and expressed in *E. coli* include the following steps:

Preparation of genomic DNA and construction of *Sbf*I genomic DNA library

Genomic DNA was prepared from *Streptomyces species Bf-61* by the 2X Kirby method ((Hopwood et al. *Genetic Manipultation of Streptomyces*. A Laboratory Manual. John Innes Foundation, Norwich. p. 77 (1985)).

Partially digested genomic DNA preparations were ligated to a digested, CIP-treated pUC19 vector into which two *Sbf*I sites had been previously engineered. The ligated DNA mixtures were used to transform *E. coli.* Transformants from each library were pooled and amplified, and plasmid DNA was prepared to generate primary plasmid libraries.

Methylase-selection

The primary plasmid libraries were challenged by digestion with *Sbf*I. These DNA digests were then transformed back into *E. coli* and plasmid DNA was prepared from some of these initial the survivors of each selected primary library. *Sbf*I digestion of these indicated that none were resistant to digestion, suggesting that none carried the *sbf*IM gene. Remaining surviving colonies from some of the challenged primary libraries were also pooled to form a secondary library, and challenged a second time with *Sbf*I. Plasmid DNA of these survivors again showed no resistance to *Sbf*I endonuclease digestion.

The non-recombinant *Sbf*I endonuclease was purified and the N-terminal amino acid sequence of this protein was determined along with the amino acid sequence from cyanobromide digested R.*Sbf*I fragments. Based on these amino acid sequences, converging sets of degenerate and non-degenerate primers were synthesized. These were used to prime PCR reactions on *Sbf*I chromosomal DNA generating DNA fragments containing the 5′ end of the *sbfIR* gene. DNA sequencing revealed an open reading frame (ORF) of 555 bp that had extensive homology to the R.*Pst*I endonuclease and to its isoschizomers, *Bsu*BI and *Xph*I. The *Pst*I recognition sequence, 5′-CTGCA/G-3′, is encompassed by the *Sbf*I recognition sequence, 5′-CCTGCA/GG-3′. These similarities in recognition sequences and amino acid sequences strongly suggested that the 555 bp ORF comprised the 5′ end of *sbfIR* gene.

Inverse PCR amplification of DNA downstream of the 5′ end *Sbf*I endonuclease gene

Following cloning and identification of the N-terminal portion of the *sbfIR* endonuclease gene, efforts were made to clone 3′ end of *sbfIR* and the adjacent downstream DNA.

Genomic DNA was digested with multiple restriction enzymes and self-ligated. The resulting circular DNA molecules were used as templates for inverse PCR. The DNA sequence at

the N-terminus of the *sbfIR* gene was used to design primers for the inverse PCR of *Sbf*I chromosomal DNA.

Not all inverse PCR reactions of self-ligated genomic DNA generated inverse PCR fragments. Only the *Hinc*II *and Hpy*CH4HIV templates produced PCR fragments that could be purified for cloning. The *Hinc*II- *and Hpy*CH4HIV- inverse PCR fragments were ligated into pUC19 and transformed into *E. coli.* Clones with PCR inserts were sequenced directly with pUC19 universal primers.

From this DNA sequence near the newly found *Hinc*II site, a new primer was designed and was used in a PCR reaction with a primer from the N-terminal *Sbf*I endonuclease sequence to generate a fragment that linked the N-terminal sequence to the inverse PCR sequence downstream of 5' *sbfIR*. The PCR fragment was purified, digested and ligated into pUC19. This ligated DNA was transformed into *E. coli and c*lones with PCR inserts were sequenced directly with pUC19 universal primers.

The *Hinc*II and *Hpy*CH4HIV PCR overlapping fragments generated ~900 bp, and ~500 bp of new sequences, respectively. Combining these sequences with the 5' *sbfIR* sequence, a complete ORF of 969 bp was found, most likely representing the *sbfIR* restriction gene. An additional ~700 bp of sequenced *Sbf*I chromosomal DNA that was downstream of *sbfIR* was compared to the known gene products in GenBank using BLAST and did not appear to contain the *Sbf*I methylase gene.

Inverse PCR amplification of DNA upstream of *SbfI* endonuclease and identification of *SbfI* methylase

Unable to isolate the *sbfIM* gene by methylase selection, combined with the inability to identify the *sbfIM* gene downstream of the *SbfI* endonuclease, efforts were made to clone DNA upstream of the *sbfIR* gene.

PCR reactions were done on the *SbfI* primary pUC19 libraries used in the initial methylase selection with two converging primers within the *sbfIR* gene. PCR showed that the *SbfI* primary libraries contained at least a portion of the *sbfIR* gene. To identify any larger upstream DNA fragments that might be contained in the initial primary libraries, a second PCR was done on the libraries with pUC19 universal primers and a primer designed from within the *sbfIR* gene oriented toward the upstream or 5′ end of *sbfIR*. These atypical PCR reactions generated fragments from the *BglII*, *BclI* and *Sau*3AI primary libraries that were larger than 1.6 kb and potentially large enough to contain upstream DNA for the *sbfIM* gene. These library generated PCR fragments were purified and DNA sequenced directly using the PCR primers.

The only readable sequence came from the *SbfI* *Sau*3AI library PCR fragment. This DNA sequence was used to design a set of new PCR primers to be used in PCR with a converging primer from within the *sbfIR* gene to hopefully directly clone the DNA upstream relative to the *sbfIR* gene from *SbfI* chromosomal

DNA. These PCR fragments were purified, digested, and ligated into pUC19 with compatible ends, followed by transformation into *E. coli*. Colony PCR identified PCR inserts and clones containing about the correct size fragment were further purified by CsCl purification and sequenced with pUC19 universal and custom primers.

DNA sequencing revealed a new upstream ORF, which was determined to be the 3′ end of the *sbfIM* gene, which was coupled and arranged in head-to-tail fashion to the *sbfIR* gene. The *sbfIM* transcriptional (TGA) stop codon was found to overlap the GTG start codon for the *Sbf*I endonulease gene. DNA sequencing across the junction revealed that along with the GTG start for the *sbfIR* gene, an additional translated amino acid, serine (S), is present making R.*Sbf*I one amino acid longer than predicted from N-terminal amino acid sequence. In fact R.*Sbf*I is actually 323 amino acids in length, not 322 amino acids. With the additional serine, the full *sbfIR* gene is 971 bp. (Figure 3).

Unpredictably during cloning, due to a designed *Hind*III site in the PCR primer, DNA sequencing revealed an unknown *Hind*III site contained within the *sbfIM* gene. This truncated the PCR fragment during cloning and in effect cut off the 5′ end of the *sbfIM* gene. Efforts were then made to clone 5′ end of *sbfIM* using inverse PCR with new primers designed from within the truncated *sbfIM* gene. Inverse PCR was performed on genomic DNA digested with *Hinc*II and self-ligated. The resulting PCR fragments were purified, ligated into pCAB16 and transformed

into *E. coli*. Plasmid DNAs were purified and sequenced. Combining this DNA sequence with the truncated 3′ *sbfIM* gene DNA sequence, the full length *sbfIM* gene of 1460 bp was revealed, which encodes a translated *Sbf*I methylase of 496 amino acids (Figure 2).

Isolation of intact *SbfIR* gene in *E.coli*

Since the *sbfIM* gene was not isolated initially, the cloning/expression strategy was to use M.*Pst*I to pre-modify *E. coli* by expressing the *pstIM* gene in a low copy-number plasmid, pACYC184, and the *sbfIR* gene in either a high copy-number constitutive vector pRRS, or a medium-copy-number, regulated vector, pLT7K.

The *sbfIR* gene was amplified from genomic DNA by PCR with Deep Vent® DNA polymerase. Following purification and digestion the PCR fragment was ligated into pRRS and pLT7K, respectively, digested with the same enzymes to create compatible ends. The pRRS-SbfIR ligation was transformed into M.*Pst*I pre-modified *E. coli* and plasmids were purified and screened for inserts. No pRRS-SbfIR clones were found. The pLT7K-SbfIR ligation was transformed into *E. coli* ER2502 (lacking T7 RNA polymerase). Positive clones were identified and these were then transferred into a M.*Pst*I *E. coli* ER2744 (containing the T7 RNA polymerase). Cell cultures were made from individual pLT7K-SbfIR transformants and induced with IPTG. Cell extracts

were prepared and assayed for *Sbf*I endonuclease activity and none produced detectable *Sbf*I endonuclease.

In another attempt, a larger PCR fragment containing the *sbfIR* gene plus 600 bp of downstream DNA (~1500 bp), was digested and ligated into pLT7K and transformed into *E. coli*. Positive clones were identified and then transferred into M.*Pst*I pre-modified *E. coli* ER2744 and induced with IPTG, assayed for *Sbf*I activity, yielding no detectable activity. This same pLT7K-SbfIR ligation of the *sbfIR* gene, including 600 bp of downstream DNA, was also directly transformed into M.*Pst*I pre-modified *E. coli* ER2744. Positive pLT7K-SbfIR clones were identified by PCR, cultured, induced with IPTG and then assayed again yielding no detectable *Sbf*I activity. This negative result indicated there was selection pressure to isolate endonuclease mutants.

In order to isolate a clone containing the *sbfIR* gene with the correct DNA sequence from the native *Streptomyces species Bf-61* strain, the *sbfIR* gene PCR was blunt-end ligated into pCAB16 at the *Bsa*AI site followed by tranformation into M.*Pst*I pre-modified *E. coli*. Clones were found to carry the *sbfIR* gene fragment, and cultured, and assayed for *Sbf*I activity on λ DNA. The extracts generated no *Sbf*I digestion pattern, however, DNA sequence from these clones showed an intact *sbfIR* gene in the opposite orientation to Plac and *mspIR*.

Expression of *SbfIR* gene in *E.coli*

The *sbfIR* gene was gel purified from the pCAB16-SbfIR plasmid, ligated into pLT7K and transformed into M.*Pst*I pre-modified *E. coli*. Clones were found to carry the *sbfIR* gene and assayed for *Sbf*I activity on λ DNA. The extracts generated partial *Sbf*I digestion pattern. DNA sequencing of pLT7K-SbfIR clones showed that each carried an intact *sbfIR* gene. Since the DNA sequence was correct for the pLT7K-SbfIR #12 plasmid, when the *sbfIM* gene was identified and cloned into pACYC184, the pLT7K-SbfIR #12 clone was then transformed into M.*Sbf*I pre-modified *E. coli*. All transformants were cultured, induced with IPTG and then assayed for *Sbf*I activity on λ DNA yielding $\sim 10^5$ units/g of wet cells from the over-producing strain.

The present invention is further illustrated by the following Examples. These Examples are provided to aid in the understanding of the invention and are not construed as a limitation thereof.

All references cited above and below are herein incorporated by reference.

EXAMPLE 1

Cloning of *SbfI* restriction-modification system in *E. coli*

1. Preparation of genomic DNA

Genomic DNA was prepared from 2g of *Streptomyces species Bf-61*, by the following steps:

a. Cell wall digestion by addition of lysozyme (2 mg/ml final), sucrose (1% final), and 50 mM Tris-HCl, pH 8.0.

b. Cell lysis by addition of 8 ml of 2X Kirby mixture: (2g Sodium tri-isopropylnaphthalene sulphonate, 12g 4-amino-salicylate, 10ml 1M Tris-HCl pH8, 6 ml phenol saturated with 50mM Tris-HCl pH 8.0 made up in 100 ml dH_20) Vortex 1 minute.

c. Removal of proteins by phenol-$CHCl_3$ extraction of DNA 2 times (equal volume).

d. Dialysis in 4 liters of TE buffer, buffer change four times.

e. RNase A treatment to remove RNA.

f. Genomic DNA precipitation in 0.4M NaCl and 0.55 volume of 100% isopropanol, spooled, dried and resuspended in TE buffer.

2. Restriction digestion of genomic DNA and construction of genomic DNA library

Varying units of restriction enzymes *Sau*3AI, *Bcl*I, *Bst*YI, *Bgl*II, *Xba*I and *Nhe*I were used to digest 10 mg genomic DNA to

achieve complete and limited partial digestion. The digested DNA was purified via phenol-$CHCl_3$ extraction and isopropanol precipitation. The *Sau*3AI-, *Bcl*I-, *Bst*YI-, and *Bgl*II-, digested DNAs were ligated to *Bam*HI- digested and CIP-treated pUC19 vector containing 2 *Sbf*I sites. *Xba*I- and *Nhe*I-, digested DNAs were ligated to the same vector *Xba*I- digested and CIP treated. Following overnight ligation, the DNA was used to transform an $endA^-$ host (ER2502, ER2683 New England Biolabs' collection (Beverly, MA)), made competent by $CaCl_2$ method. Approximately 2-5,000 Ap^R transformants were obtained from each library. For each enzyme, colonies were pooled and amplified in 500ml LB + Ap overnight. Plasmid DNA was prepared by CsCl gradient purification, resulting in a primary library.

3. Attempt to Clone *SbfIM* gene by methylase selection

The primary plasmid DNA library (1 mg DNA) was challenged by digestion with ~30 units of *Sbf*I at 37°C for 1 hour. The digested DNA was transferred into ER2502 or ER2683 by transformation, resulting in ~750 Ap^R survivors from all libraries. Plasmid DNA from ~120 survivors was prepared by the Compass Mini Plasmid Kit method, followed by *Sbf*I digestion. No resistant clones were found from any of the libraries. Some remaining survivors (*Sau*3AI, *Bst*YI, *Xba*I and *Nhe*I) were also pooled separately to form secondary libraries, challenged with *Sbf*I a second time, followed by the same survivor Plasmid DNA purification, again no resistant clones were found.

4. Identification of the SbfI endonuclease

The non-recombinant *Sbf*I endonuclease was purified to near homogeneity and the purified protein was subjected to SDS-PAGE. A protein band of ~36 kDa was detected. The N-terminal amino acid sequence was determined as (MN)SDGIDGTV ASIDTARALLKRFGFDAQRYNV (SEQ ID NO:5). The 36 kDa protein was digested with cyanogen bromide and a 4.5 kDa fragment amino acid sequence was determined as (M)VEEFVPRFAPRSTV LYLGDTRGKHSLFEEEI (SEQ ID NO:6). Using these two amino acid sequences, converging sets of degenerate and non-degenerate primers were designed to PCR the beginning of the *sbfIR* gene from chromosomal DNA.

Two sets primers were synthesized with the following sequences:

5'atgaactccgacggcatcgac3' sbfN-1
(SEQ ID NO:7)

5'aanacyaartcnaccat3' sbf45a
(SEQ ID NO:8)

5'atgaacagcgacggcatcgac3' sbfN-1b
(SEQ ID NO:9)

5'aasaccaactcctcsaccat3' sbf45a-2
(SEQ ID NO:10)

The primers were used in two separate PCR reactions: sbfN-1 + sbf45a, sbfN-1b + sbf45a-2. PCR conditions were 95°C

for 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C 1 min for 25 cycles with Deep Vent® DNA polymerase. The PCR of *Sbf*I chromosomal DNA with the above primers generated ~550 bp DNA fragment. The PCR fragments were gel purified, phenol-CH_3Cl extracted and isopropanol precipitated. The resuspended PCR fragment was blunt-end ligated into pCAB16 at the *Bsa*AI site followed by transformation into *E.coli* ER2502. pCAB16 clones with PCR inserts were sequenced using the following sequencing primers:

5'ggagccatacagagagcgatttattcg3' 167
(SEQ ID NO:11)

5'ttgaaatcgaattaataagtctggatg3' 168
(SEQ ID NO:12)

DNA sequencing identified an open reading frame (ORF) of 555 bp DNA fragment containing the 5' end of the *sbfIR* gene.

5. Inverse PCR amplification of DNA downstream of the 5' end *SbfI* endonuclease gene

After identification of the N-terminus of the endonuclease gene, efforts were made to clone adjacent downstream DNA. DNA sequence at the N-terminus of the *sbfIR* gene was used as the template for primer design.

Four primers were synthesized:

5'ccagtccatgatcttctgaacgcc3' 5B
(SEQ ID NO:13)

5'cttcggcagtgggttgataatggc3' 3B
(SEQ ID NO:14)

5'agggagatcgacagagatcatcgc3' 5C
(SEQ ID NO:15)

5'tactgcgcgggcgctgctaaagcg3' 3A
(SEQ ID NO:16)

As a positive control, converging primers 3A and 5C, were used on *Sbf*I chromosomal DNA. For the inverse PCR, genomic DNA was individually digested with *Bst*BI, *Bst*UI, *Dra*I, *Hinc*II, *Hpy*CH4IV, *Rsa*I and *Sca*I. The digestions were inactivated at 65°C for 20 min. Self-ligation was set up at a low DNA concentration at 2 mg/ml overnight at 17°C. The resulting circular DNA preps were used as the templates for inverse PCR. PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 62°C for 1 min, 72°C for 2 min for 25 cycles. Converging primers 3A and 5C generated the *sbfIR* ~400 bp control fragment. Inverse PCR products were found in the *Hinc*II and *Hpy*CH4IV templates. The PCR products were gel-purified, phenol/CH_3Cl-extracted and isopropanol-precipitated. Immediately downstream of the 3B primer within the *sbfIR* N-terminus is an *Apo*I site. This *Apo*I site was used to digest the inverse PCR products at this site, followed by overnight ligation into *Eco*RI- and *Hinc*II- digested pUC19. The ligated DNA was transferred into ER2502 and ER2683 by transformation. Plasmids were identified that contained the inverse PCR fragment and sequenced directly with pUC19 universal primers 1233 and 1224. Using this DNA sequence, another direct PCR was done with 3B and a newly designed converging primer Sb-3 having the following sequence:

5'gcggcaaccttcatccgg3' Sb-3
(SEQ ID NO:17)

PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C for 2 min for 25 cycles. A PCR fragment containing ~1200 bp of the C-terminal end of *sbfIR* and downstream DNA was purified by phenol/CH3Cl -extraction and isopropanol precipitation. The PCR fragment was digested with *Apo*I and ligated into *Eco*RI- and *Hinc*II- digested pUC19. The ligated DNA was transformed into ER2683 and plasmid DNA was purified. Plasmids containing the PCR fragment were sequenced with pUC19 universal primers 1233 and 1224. After cloning, then sequencing, an ORF with 969 bp long was found. This gene was judged likely to be the *sbfIR* restriction gene coding for *Sbf*I endonuclease as predicted by N-terminal R.*Sbf*I amino acid protein sequence. An additional ~700 bp of sequenced *Sbf*I chromosomal DNA that was downstream of *sbfIR* did not appear to contain the *Sbf*I methylase gene.

6. PCR amplification of DNA upstream of *SbfI* endonuclease

After failing to identify the *sbfIM* gene downstream of the endonuclease gene, efforts were made to clone adjacent DNA upstream to *sbfIR*. A PCR reaction was done on the original pUC19 primary libraries of *Sau*3AI, *Bcl*I, *Bst*YI, *Bgl*II, *Xba*I and *Nhe*I to check for the presence of *sbfIR* using converging primers 5C and 3A:

5'agggagatcgacagagatcatcgc3' 5C
(SEQ ID NO:18)

5'tactgcgcgggcgctgctaaagcg3' 3A
(SEQ ID NO:19)

PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 62°C for 1 min, 72°C for 2 min for 25 cycles. All libraries, except the *Xba*I library, contained the ~400 bp *sbfIR* fragment that should be generated by these two primers from *sbfIR* DNA sequence. Two separate PCR reactions were done with pUC19 universal primers 1233 and 1224 with primer 5B, in the direction of 5' end of *sbfIR*, to determine if some of these pUC19 libraries might contain DNA upstream of the sequence *sbfIR*, perhaps containing the *Sbf*I methylase gene. The libraries tested were the *Bgl*II, *Bcl*I and *Sau*3AI pUC19 libraries. The PCR primers have the following sequence:

5'agcggataacaatttcacacagga3' 1233 pUC19 Universal primer
(SEQ ID NO:20)

5'cgccagggttttcccagtcacgac3' 1224 pUC19 Universal primer
(SEQ ID NO:21)

5'ccagtccatgatcttctgaacgcc3' 5B
(SEQ ID NO:22)

PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 62°C for 1 min, 72°C for 8 min for 25 cycles. The 1233 and 5B PCR primers generated a predominant fragment of ~1.6 kb for *Bgl*II and *Sau*3AI libraries; and ~4.0 kb fragment from the BclI library, both perhaps large enough to contain the whole *Sbf*I methylase gene (*sbfIM*). The PCR products were gel-purified,

phenol-CH_3Cl extracted and isopropanol precipitated, followed by direct sequencing of the PCR products. Only the PCR fragment from the *SbfI* *Sau*3AI-pUC19 partial library generated DNA sequence sufficiently readable to design new PCR primers for a direct PCR of the upstream region in its entirety. The *Sau*3AI library DNA sequence upstream of the *sbfIR* gene was used as the template for primer design and a new converging primer 5B-2 was made from *sbfIR* DNA sequence toward the upstream DNA sequence.

The primers were synthesized with the following sequences:

5'tggggcgaattcccagtccatgatcttctgaacgcc3' 5B-2
(underlined nt, *EcoRI* site) (SEQ ID NO:23)

5'tggggcaagcttgatcaggtccgtg3' S3-1
(underlined nt, *Hind*III site) (SEQ ID NO:24)

5'tggggcaagcttcgcctgctggttgacc3' S3-2
(underlined nt, *Hind*III site) (SEQ ID NO:25)

5'tgtggggcaagcttcgccccggtcgtcc3' S3-3
(underlined nt, *Hind*III site) (SEQ ID NO:26)

5'tggggcaagcttctgcgatccgctgcc3' S3-4
(underlined nt, *Hind*III site) (SEQ ID NO:27)

5'tggggcaagcttcgttggcggtgctcccgc3' S3-5
(underlined nt, *Hind*III site) (SEQ ID NO:28)

PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C for 1 min for 25 cycles. PCR fragments of ~1550 bp were found with both 5B-2 +S3-2 or 5B-2+S3-3

primers on *Sbf*I chromosomal DNA. The PCR fragments were gel-purified, phenol-CH_3Cl extracted and isopropanol precipitated, followed by *Eco*RI and *Hind*III digestion. The digested PCR DNA was heated to 65°C for 15 min and ligated overnight at 17°C into *Eco*RI- and *Hind*III- digested pUC19. The ligated DNA was transformed into ER2744 and colony PCR was done on 10 colonies of each with universal pUC19 primers 1233 and 1224. PCR conditions were 94°C for 1 min, 1 cycle; 94°C for 10 sec, 62°C for 1 min, 72°C for 1 min for 25 cycles. Most colonies contained a PCR fragment approximately 1500 bp. Plasmid DNA was purified for 6 clones, 3 for each PCR fragment, by CsCl method and then sequenced using primers 1233 and 1224. DNA sequence was only obtained from clone #5 which contained the PCR fragment generated by primers 5B-2 + S3-3. Clone #5 contained a *Hind*III to *Eco*RI fragment of ~1200 bp, slightly less then the original PCR fragment. The entire fragment was sequenced with seven additional primers. The sequencing primers have the following sequences:

5'gagcaatgtcacagcgctacggac3' (51)
(SEQ ID NO:29)

5'gatccaacacagtcgagttcaacc3' (52)
(SEQ ID NO:30)

5'aacggcaaacggcgaaagaggacc3' (53)
(SEQ ID NO:31)

5'cctcatgctttggttgaactcgac3' (54)
(SEQ ID NO:32)

5'tcgttgctttcagtgtgaggccgc3' (55)
(SEQ ID NO:33)

5'cgttgtgcccatggtttatcagtc3' (56)
(SEQ ID NO:34)

5'ctcttgctttaccttcgtgtccgc3' (57)
(SEQ ID NO:35)

After sequencing about 1.2 kb, a truncated ORF of 1040 bp long was found upstream of the *sbfIR* gene. This sequence most likely encoded the 3' end of the *SbfI* methylase gene. When the amino acid sequence of this ORF was compared to the known gene products in GenBank using BLAST, it showed very high homology to *N6*-methyl adenine methyltransferases especially those belonging to the *Pst*I recognition family. The *sbfIM* transcriptional (TGA) stop codon was found to overlap the (GTG) start codon for the *Sbf*I endonulease gene (*sbfIR*). DNA sequencing across the junction revealed that along with the GTG start, an additional amino acid, serine (S) is present making R.*Sbf*I one amino acid longer than predicted from N-terminal amino acid. The original incorrect sequence is (MN)SDGIDGTVAS IDTARALLKRFGFDAQRYNV (SEQ ID NO:36). In fact R.*Sbf*I is actually 323 amino acids in length, not 322 amino acids, and the N-terminal sequence is: MNSSDGIDGTVASIDTARALLKRFGFDAQ RYNV (SEQ ID NO:37). With the additional Serine, the full *Sbf*I endonuclease gene (*sbfIR*) is 971 bp. DNA sequencing of #5 also revealed an unknown *Hind*III site within *sbfIM* which shortened the original 5B-2 + S3-3 PCR fragment. DNA between the S3-3 primer to this *Hind*III site was lost during cloning.

7. Inverse PCR amplification of DNA upstream of *Sbf*I endonuclease and identification of *Sbf*I methylase

After identification of the truncated methylase gene, efforts were made to clone adjacent DNA encoding the 5' end of *sbfIM* using inverse PCR. Using this *sbfIM* DNA sequence new primers were designed having the following sequence:

5'ggccacgtaaacgttcggtacatc3' (A1)
(SEQ ID NO:38)

5'tcatttcgctcaaagagcaggggc3' (B1)
(SEQ ID NO:39)

The genomic DNA was digested with *Hinc*II in appropriate restriction buffer and inactivated at 65°C for 20 min. Self-ligation was set up at a low DNA concentration at 2 mg/ml overnight at 17°C. The circular DNA product was used as the template for inverse PCR. PCR conditions were 94°C for 5 min, 1 cycle; 94°C for 30 sec, 62°C for 1 min, 72°C for 1 min for 25 cycles. The PCR fragment was gel purified from an agarose gel, phenol/CH_3Cl-extracted and isopropanol precipitated. The resuspended PCR fragment was blunt-end ligated at 17°C overnight into pCAB16 digested at the *Bsa*AI site followed by tranformation into ER2502 *E. coli* cells. Plasmid DNA was purified from twelve colonies. Ten appeared to contain the PCR DNA fragment. Four clones were sequenced directly with the following primers:

5'ggagccatacagagagcgatttattcg3' 167
(SEQ ID NO:40)

5'ttgaaatcgaattaataagtctggatg3' 168
(SEQ ID NO:41)

DNA sequencing identified an open reading frame (ORF) of 495 bp DNA fragment containing the 5' end of *sbfIM*. The combined DNA sequence upstream of the *sbfIR* gene revealed the total *Sbf*I methylase gene (*sbfIM*). The *sbfIM* gene is 1460 bp which encodes a translated *Sbf*I methylase of 486 amino acids. Transcription of M and R genes is oriented in the same direction. They are arranged in head-to-tail fashion (Figure 1).

EXAMPLE 2

Expression of *SbfIR* gene in *E.coli*

Since the preferred methylase selection method failed to yield a *Sbf*I methylase gene, efforts were made initially to clone the *sbfIR* gene in a M.*Pst*I pre-modified *E. coli* strain in order to establish an *sbfIR* clone with the correct DNA sequence. The *pst*IM gene was first amplified from pBEA-14 plasmid containing the *Pst*I methylase (provided by Bill Jack and Lucia Greenough, New England Biolabs (Beverly, MA)). The *pstIM* gene was cloned into a low copy number plasmid pACYC184 with p15A origin and Cm^R selection marker. The PCR primers have the following sequences:

5'ttccgggatccggaggtttaaaatatgactaagcggcaacacaattacctatatctc3'
(Pst5M, underlined nt, *Bam*HI site) (SEQ ID NO:42)

5'ttgccgcatgcgtcgttacattagctgcaaactctgattgattatttc3'
(Pst3M, underlined nt, *Sph*I site) (SEQ ID NO:43)

PCR conditions were 95°C for 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C 1 min for 25 cycles with Deep Vent® DNA polymerase. The PCR product was phenol-CH_3Cl extracted and isopropanol precipitated, followed by *Bam*HI and *Sph*I digestion overnight at 37°C. Digested DNA was ligated to pACYC184 with compatible ends. Following ligation overnight, the DNA was transferred into ER2744 by transformation. After screening 10 plasmids isolated from individual Cm^R (33 mg/ml) transformants, one clone pACYC184-PstIM #4 contained the correct *pstIM* fragment as shown by *Bam*HI and *Sph*I digestion. ER2744 [pACYC184-PstIM #4] was made competent by $CaCl_2$ method. The premodified host ER2744 [pACYC184-PstIM] was used for establishing of the *SbfI* endonuclease.

Two PCR primers were synthesized for PCR amplification of the *sbfIR* gene.

5'tggctgcaggGatccggaggtttaaaatatgaacagcagtgacggcatcgacggaacggtagc3'
(sbf5R-2, underlined nt, *Bam*HI site) (SEQ ID NO:44)

5'tcgggccccgggctcgagtctaacgttcgtacggcccaagaaatctagacc3'
(sbf3RT7, underlined nt, *Xho*I site) (SEQ ID NO:45)

PCR conditions were 95°C 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C for 1 min for 25 cycles with Deep Vent® DNA polymerase. PCR DNA containing the *sbfIR* gene was amplified from genomic DNA and purified by phenol/ CH_3Cl - extraction and CH_3Cl extraction, precipitated with isopropanol, dried and resuspended in TE buffer. The PCR DNA was blunt-end

ligated to pCAB16 at the *Bsa*AI site. The ligated DNA was transformed into M.*Pst*I pre-modified *E.coli* ER2744 [pACYC184-PstIM] and Ap^R Cm^R transformants were selected at 37°C [Ap^R Cm^R :(100 mg/ml) and (33 mg/ml]. After screening 12 plasmids, 4 clones were found to contain inserts. 500 ml LB + Amp cultures of clones #1 and #7 were grown overnight at 37°C. Plasmid DNA was purified from 450 ml of each cell culture by the CsCl method, and the other 50 ml of the cells of each were harvested by centrifugation and resuspended in 2 ml sonication buffer (10 mM Tris-HCl, pH 8, 0.1 mM EDTA, 50 mM NaCl, 1 mM ß-mercaptoethanol). Cells were lysed by sonication and cell debris removed by centrifugation. Cell lysate was assayed on l DNA for *Sbf*I activity. No *Sbf*I activity was detected for either pCAB16-SbfIR #1 or #7 clones. Both pCAB16-SbfIR clones were sequenced directly with the following primers:

5'ggagccatacagagagcgatttattcg3' 167
(SEQ ID NO:46)

5'ttgaaatcgaattaataagtctggatg3' 168
(SEQ ID NO:47)

DNA sequence showed an intact *sbfIR* in the opposite orientation to Plac and *mspIR*. The pCAB16-sbfIR #1 and #7 clones contained correct DNA sequence for *sbfIR*, so an attempt was made to subclone the *Sbf*I endonuclease gene from pCAB16-sbfIR #1 and #7 into pLT7K, and then transform into ER2744 [pACYC184-PstIM]. Using the flanking *Bam*HI and *Xho*I sites designed within the PCR primers, 10 µg of each pCAB16-sbfIR plasmid was digested with *Bam*HI and *Xho*I at 37°C for 2

hours and the *sbf*IR fragment were gel purified from a agarose gel, phenol-CH_3Cl extracted and isopropanol precipitated. The resuspended PCR fragment was was ligated at 17°C overnight into pLT7K with compatible ends, followed by transformation into ER2744 [pACYC184-PstIM] and plated on Ap^R Cm^R plates at 37°C. Plasmid DNA's were purified from 18 colonies (9 each from gel pure pCAB16-sbfIR #1 and #7, respectfully), 6 clones were found to carry the PCR insert. pLT7K-SbfIR #5, #12 and #14 were inoculated into pre-warmed 10 ml cultures containing LB+Ap^R Cm^R and grown at 37°C overnight without shaking. 2 ml of the overnight cultures were diluted in pre-warmed 50 ml cultures containing LB+Ap^R Cm^R and grown at 37°C to an OD590 of between 0.8 and 1.0, IPTG was to added to 85 mg/L and induced at 30°C for ~2 hours. Cells were harvested and lysed by sonication. Clarified cell lysates were assayed for *Sbf*I activity on l DNA. The extracts generated partial *Sbf*I digestion pattern. pLT7K-SbfIR #12 was sequenced with the following primers:

5'tactgcgcgggcgctgctaaagcg3' 3A
(SEQ ID NO:48)

5'aatttctgctctcgcctgccgggc3' 3C
(SEQ ID NO:49)

5'ccagtccatgatcttctgaacgcc3' 5B
(SEQ ID NO:50)

5'agggagatcgacagagatcatcgc3' 5C
(SEQ ID NO:51)

The DNA sequence of pLT7K-SbfIR #12 showed that it carried an intact *sbfIR* gene. Shortly after this result, the *Sbf*I

methylase gene (*sbfIM*) was completely identified by inverse PCR and sequenced. A final strategy was employed in which the *sbfIM* gene was expressed from a low-copy-number plasmid and the endonuclease gene from pLT7K-SbfIR #12 was then transferred into this *Sbf*I pre-modified *E. coli* host. The *sbfIM* gene was first amplified from genomic DNA in a PCR reaction and was cloned into a low copy number plasmid pACYC184 with p15A origin and Cm^R selection marker. The PCR primers have the following sequences:

5'tggccgggatccggaggtttaaaatatgcatccgatcgccagcactgaaactcg ccgc3'
(sbf5M, underlined nt, *Bam*HI site) (SEQ ID NO:52)

5'ttgccgcatgcctcacgaggcagattccggaatctcacagaagagtc3'
(sbf3M, underlined nt, *Sph*I site) (SEQ ID NO:53)

PCR conditions were 95°C 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C for 1 min for 30 cycles with Deep Vent® DNA polymerase. PCR DNA containing the *sbfIM* gene was amplified from genomic DNA and purified by phenol/CH3Cl - extraction, precipitated with isopropanol, dried and resuspended in TE buffer. The PCR DNA was blunt-end ligated to pCAB16 at the *Bsa*AI site. The ligated DNA was transformed into ER2502 and Cm^R transformants were selected at 37°C. After screening 12 plasmids, two pCAB16-sbfIM clones, #3 and #5 were completely sequenced with the following primers:

5'ggagccatacagagagcgatttattcg3' 167
(SEQ ID NO:54)

5'ttgaaatcgaattaataagtctggatg3' 168

(SEQ ID NO:55)

5'ctttcccgtacttacaccgatgcc3' sbf-M1
(SEQ ID NO:56)

5'tcctcgatcttgggcgaacgggcc3' sbf-M2
(SEQ ID NO:57)

DNA sequence showed an intact *sbfIM* in both orientations to Plac and *mspIR*. The pCAB16-sbfIR #3 and #5 clones contained correct DNA sequence for the *sbfIM* gene, so an attempt was made to subclone the *Sbf*I methylase gene from pCAB16-sbfIR #3 and #7 into the low copy number plasmid pACYC184 with p15A origin and Cm^R selection marker. Using the flanking *Bam*HI and *Sph*I sites designed within the PCR primers, 5 µg of each pCAB16-SbfIM plasmid was digested with *Bam*HI and *Sph*I at 37°C for 2 hours and the *sbf*IM fragment were gel purified from a agarose gel and combined, phenol-CH_3Cl extracted and isopropanol precipitated. The re-suspended PCR fragment was was ligated at 17°C overnight into pACYC184 with compatible ends, followed by transformation into ER2848 and then selected for Cm^R transformants. Plasmid DNA was purified from 8 colonies and PCR was done on two clones, #6 and #7, with primers sbf5M and sbf3M. PCR conditions were 95°C 5 min, 1 cycle; 95°C for 1 min, 54°C for 1 min, 72°C for 1 min for 25 cycles with Deep Vent® DNA polymerase. pACYC184-SbfIM #6 and #7 contained the correct size insert for *sbfIM*.

The plasmid pACYC184-SbfIM #7 was transferred into ER2848 to premodify *E. coli*. Competent cells were made by $CaCl_2$

method and the final strategy was employed in which the *sbfIM* gene was expressed from a low-copy-number plasmid [pACYC184-SbfIM] and the endonuclease gene from [pLT7K-SbfIR]. Isolate pLT7K-SbfIR #12 was transferred into ER2848 [pACYC184-SbfIM] and plated on $Ap^R Cm^R$ plates at 37°C overnight. Two individual colonies were inoculated into 10 ml LB+$Ap^R Cm^R$ and grown at 37°C overnight. 1 ml of each overnight culture was inoculated into 50 ml of LB+$Ap^R Cm^R$ and grown at 37°C to OD590 0.8 to 1.0, then the culture temperature was then lowered to 30°C, followed by IPTG (85 mg/L) induction at 30°C for 2 hours to overnight. Both individual clones expressed R.*SbfI* at more ~10^5 u/g per gram of wet *E. coli* cells (Figure 8).

The strain NEB#1500, ER2848 [pACYC184-SbfIM #7, pLT7K-SbfIR #12 has been deposited under the terms and conditions of the Budapest Treaty with the American Type Culture Collection on August 5, 2003 and received ATCC Accession No. PTA-5371.